Processing Bulleted Lists

ASCII English free text should ideally have some semblance of normal English syntactic structure. Problems may occur, however, when MetaMap processes text consisting of nothing but a long sequences of nouns, which can take a long time to process and produce suboptimal results. Be especially careful of running MetaMap on the results of screen-scraping bulleted lists. For example, the text below appears as a bulleted list in the *The Technology Being Reviewed* section of PubMed Central (PMC3379165):

The following energy-based techniques were considered for assessment:

- transurethral electrovaporization of the prostate (TUVP)
- transurethral electrovapor resection of the prostate (TUVRP)
- transurethral electrovaporization of the prostate using bipolar energy (plasmakinetic vaporization of the prostate [PKVP])
- plasmakinetic vaporization of the prostate (PKVP)
- visual laser ablation of the prostate (VLAP)
- transurethral ultrasound guided laser incision prostatectomy (TULIP)
- contact laser vaporization of the prostate (CLV)
- interstitial laser coagulation (ILC)
- holmium laser resection of the prostate (HoLRP)
- holmium laser enucleation of the prostate (HoLEP)
- holmium laser ablation of the prostate (HoLAP)
- potassium titanyl phosphate (KTP) laser
- transurethral microwave thermotherapy (TUMT)
- transurethral needle ablation (TUNA)

but in PubMed itself (PMID 23074487) it appears instead as

THE TECHNOLOGY BEING REVIEWED: The following energy-based techniques were considered for assessment: transurethral electrovaporization of the prostate (TUVP)transurethral electrovapor resection of the prostate (TUVRP)transurethral electrovaporization of the prostate using bipolar energy (plasmakinetic vaporization of the prostate [PKVP])visual laser ablation of the prostate (VLAP)transurethral ultrasound guided laser incision prostatectomy (TULIP)contact laser vaporization of the prostate (CLV)interstitial laser coagulation (ILC)holmium laser resection of the prostate (HoLRP)holmium laser enucleation of the prostate (HoLEP)holmium laser ablation of the prostate (HoLAP)potassium titanyl phosphate (KTP) lasertransurethral microwave thermotherapy (TUMT)transurethral needle ablation (TUNA)

which would cause MetaMap some difficulty because the text contains no syntactic structure.

A measure of last resort: Hand-edit MetaMap's input text by inserting a semicolon (":") where the original bullets had been.